

The replica symmetric region in the Sherrington-Kirkpatrick mean field spin glass model. The Almeida-Thouless line

Francesco Guerra*

Dipartimento di Fisica, Università di Roma “La Sapienza”

INFN, Sezione di Roma1, Piazzale A. Moro 2, 00185 Roma, Italy

February 6, 2008

Abstract

In previous work, we have developed a simple method to study the behavior of the Sherrington-Kirkpatrick mean field spin glass model for high temperatures, or equivalently for high external fields. The basic idea was to couple two different replicas with a quadratic term, trying to push out the two replica overlap from its replica symmetric value. In the case of zero external field, our results reproduced the well known validity of the annealed approximation, up to the known critical value for the temperature. In the case of nontrivial external field, our method could prove the validity of the Sherrington-Kirkpatrick replica symmetric solution up to a line, which fell short of the Almeida-Thouless line, associated to the onset of the spontaneous replica symmetry breaking, in the Parisi Ansatz. Here, we make a strategic improvement of the method, by modifying the flow equations, with respect to the parameters of the model. We exploit also previous results on the overlap fluctuations in the replica symmetric region. As a result, we give a simple proof that replica symmetry holds up to the critical Almeida-Thouless line, as expected on physical grounds. Our results are compared with the characterization of the replica symmetry breaking line previously given by Talagrand. We outline also a possible extension of our methods to the broken replica symmetry region.

* e-mail: francesco.guerra@roma1.infn.it

1 Introduction

It is very well known, on physical grounds, that for the mean field spin glass model, introduced by Sherrington and Kirkpatrick in [1], the replica symmetric solution holds in a region of the parameters characterized by high temperature and/or high external magnetic field, as shown for example in [2], and references quoted there. Recently, Michel Talagrand [3] has been able to give a complete characterization of this region, through a method which exploits the broken replica symmetry bounds, established in [4], extended to the case of two coupled replicas. Talagrand method is remarkable, not only because it gives a full characterization of the replica symmetric region, but also because it can be easily extended, with some technical effort, to the broken replica symmetry region. In this way, it has been possible to establish the rigorous validity of the Parisi Ansatz for the free energy density in the infinite volume limit [5].

The motivation of this paper can be easily explained. It is surely true that some features of the broken replica symmetry phenomenon can be found also in the region of the parameters where replica symmetry still holds. This is a general aspect of statistical physics. Let us consider for example a system in its liquid phase near to the onset of crystallization. While the liquid phase is translation and rotation invariant, one can see features of the crystal phase, where translation and rotation invariance are broken, through a study of the correlations, near the transition line. In the approach of Talagrand, the broken replica symmetry features, for coupled replicas, are exploited in an essential way in order to characterize the boundaries of the replica symmetric region. However, it should be possible, in principle, to characterize the replica symmetric region in terms of its own physical content, without resorting to features of the broken replica symmetry phenomenon. We will show that this is possible indeed.

Our method is very simple. It can be considered a refinement of the interpolation method introduced in [6]. We will exploit also in a crucial way the established fact that overlap fluctuations are Gaussian in the replica symmetric region, as shown for example in [7] and [3].

The paper is organized as follows. In section 2 we recall the main features of our improved interpolation method. We consider an abstract general setting, and derive upper and lower bounds for the generalized free energy. These bounds are the essential ingredients to be exploited for the rigorous study of mean field spin glass models. In section 3 we introduce the mean field spin glass model in the replica symmetric region, and state the main results. The quadratic replica coupling for two replicas is introduced in section 4, and the main features of the corresponding replica symmetric region are

investigated. Section 5 is devoted to the calculation of overlap fluctuations for systems containing two coupled replicas. These calculations follow the pattern outlined in [8] in the case of a single replica, and are a very simple extension of the treatment given in [7]. The main results of this paper are proven in section 6, where it is shown in particular that the replica symmetric region extends up to the Almeida-Thouless line, as expected. We make also the comparison with the replica symmetric region given by Talagrand. In particular, we are able to isolate, in full generality, the parameter ruling the validity of a generic Ansatz for the free energy density. Section 7 is dedicated to some conclusion and outlook for future developments.

2 Upper and lower bounds from interpolation methods

For the sake of future reference, we give here a complete and self-contained treatment. Let us introduce the following general setting. Let U_i be a family of centered Gaussian random variables, $i = 1, \dots, K$, with covariance matrix given by $E(U_i U_j) \equiv S_{ij}$. We treat the set of indices i now as configuration space for some statistical mechanics system, with partition function Z and quenched free energy given by

$$E \log \sum_i w_i \exp(\sqrt{t} U_i) \equiv E \log Z, \quad (1)$$

where $w_i \geq 0$ are generic weights, and t is a parameter ruling the strength of the interaction.

The following well known derivation formula plays a leading role in the interpolation method

$$\begin{aligned} \frac{d}{dt} E \log \sum_i w_i \exp(\sqrt{t} U_i) = \\ \frac{1}{2} E(Z^{-1} \sum_i w_i \exp(\sqrt{t} U_i) S_{ii} \\ - \frac{1}{2} E(Z^{-2} \sum_i \sum_j w_i w_j \exp(\sqrt{t} U_i) \exp(\sqrt{t} U_j) S_{ij}). \end{aligned} \quad (2)$$

The proof is straightforward. Firstly we perform directly the t derivative. Then, we notice that the random variables appear in expressions of the form $E(U_i F)$, where F are functions of the U 's. These can be easily handled through

the following integration by parts formula for generic Gaussian random variables, analogous to the Wick theorem in quantum field theory,

$$E(U_i F) = \sum_j S_{ij} E\left(\frac{\partial}{\partial U_j} F\right). \quad (3)$$

Therefore, we see that always two derivatives are involved. The two terms in (2) come from the action of the U_j derivatives, the first acting on the Boltzmann factor, and giving rise to a Kronecker δ_{ij} , the second acting on Z^{-1} , and giving rise to the minus sign and the duplication of variables.

The derivation formula can be expressed in a more compact form by introducing replicas and suitable averages. In fact, let us introduce the state ω acting on functions F of i as follows

$$\omega(F(i)) = Z^{-1} \sum_i w_i \exp(\sqrt{t} U_i) F(i), \quad (4)$$

together with the associated product state Ω acting on replicated configuration spaces i_1, i_2, \dots, i_s . By performing also a global E average, finally we define the averages

$$\langle F \rangle_t \equiv E \Omega(F), \quad (5)$$

where the subscript is introduced in order to recall the t -dependence of these averages.

Then, the equation (2) can be written in a more compact form

$$\frac{d}{dt} E \log \sum_i w_i \exp(\sqrt{t} U_i) = \frac{1}{2} \langle S_{i_1 i_1} \rangle_t - \frac{1}{2} \langle S_{i_1 i_2} \rangle_t. \quad (6)$$

By following the same procedure, we can also easily establish the following general derivation formula for $\langle \rangle$ averages

$$\frac{d}{dt} \langle F(i_1) \rangle_t = \frac{1}{2} \langle F_{i_1} (S_{i_1 i_1} - S_{i_2 i_2} - 2S_{i_1 i_2} + 2S_{i_1 i_3}) \rangle_t. \quad (7)$$

Here, up to three replicas are involved. In fact, the first t derivative will let a second replica appear, while the subsequent $\partial/\partial A_j$ derivatives will introduce also a third replica.

Let us now go to the interpolation and comparison arguments. Let U_i and \hat{U}_i , for $i = 1, \dots, K$, be independent families of centered Gaussian random variables, with covariance matrices given by $E(U_i U_j) \equiv S_{ij}$ and $E(\hat{U}_i \hat{U}_j) \equiv \hat{S}_{ij}$. For $0 \leq t \leq 1$, introduce the interpolating function

$$\alpha(t) \equiv E \log \sum_i w_i \exp(\sqrt{t} U_i) \exp(\sqrt{1-t} \hat{U}_i), \quad (8)$$

so that we have the boundary terms

$$\alpha(0) \equiv E \log \sum_i w_i \exp(\hat{U}_i) \quad (9)$$

$$\alpha(1) \equiv E \log \sum_i w_i \exp(U_i). \quad (10)$$

By applying the t derivative as before, and taking into account that the $1-t$ term will bring an extra minus sign, we arrive to the basic interpolation formula

$$\frac{d}{dt}\alpha(t) = \frac{1}{2}\langle S_{i_1 i_1} - \hat{S}_{i_1 i_1} \rangle_t - \frac{1}{2}\langle S_{i_1 i_2} - \hat{S}_{i_1 i_2} \rangle_t. \quad (11)$$

In the simplest case, the comparison argument runs as follows. Assume that the covariances satisfy the inequalities for generic configurations

$$E(U_i U_j) \equiv S_{ij} \geq E(\hat{U}_i \hat{U}_j) \equiv \hat{S}_{ij}, \quad (12)$$

and the equalities along the diagonal

$$E(U_i U_i) \equiv S_{ii} = E(\hat{U}_i \hat{U}_i) \equiv \hat{S}_{ii}, \quad (13)$$

then we have

$$\frac{d}{dt}\alpha(t) \leq 0, \quad (14)$$

with the obvious consequences

$$\alpha(1) \leq \alpha(0), \quad \alpha(t) \leq \alpha(0). \quad (15)$$

In the applications, $\alpha(1)$ is the quantity of interest, while $\alpha(0)$ is connected to some comparison system, and can be easily calculated. Therefore, we call (15) an upper bound formula.

Considerations of this kind are present in the mathematical literature of some years ago. Two typical references are [9] and [10].

If the stated conditions for the covariances (12,13) hold, then we also have

$$\frac{d}{dt}(\alpha(0) - \alpha(t)) = \frac{1}{2}\langle \Delta_{12} \rangle_t, \quad (16)$$

where

$$\Delta_{12} \equiv S_{i_1 i_2} - \hat{S}_{i_1 i_2} \geq 0. \quad (17)$$

By integration, we have also

$$\alpha(0) = \alpha(1) + \frac{1}{2} \int_0^1 \langle \Delta_{12} \rangle_t dt. \quad (18)$$

We call “sum rule” an expression of this kind, by following the terminology introduced in [8].

The derivation formula and the comparison argument are not restricted to the Gaussian case. Generalizations in many directions are possible. For the diluted spin glass models and optimization problems we refer for example to [11], and to [12], and references quoted there.

Let us show how to derive lower bounds, in the opposite direction. We follow the general strategy expounded in [6], with some relevant improvement.

Introduce a system made by two replicas with some appropriate coupling between them. Now the interpolating quenched free energy is defined as

$$\begin{aligned} \tilde{\alpha}(\lambda, t) \equiv \\ \frac{1}{2} E \log \sum_{ij} w_i w_j \exp(\sqrt{t}(U_i + U_j) + \sqrt{1-t}(\hat{U}_i + \hat{U}_j)) \exp(\lambda(S_{ij} - \hat{S}_{ij})), \end{aligned} \quad (19)$$

depending on two parameters, the interpolating parameter $0 \leq t \leq 1$, and the coupling parameter $\lambda \geq 0$. We have also introduced the factor $1/2$ in order to keep the same overall normalization. In fact, we have $\tilde{\alpha}(0, t) = \alpha(t)$, if λ is put to zero.

It is simple to calculate the t derivative by a direct application of the general formula (11). Of course, now we have to take into account that the original system is made of two (coupled) replicas, let us say $(1, 2)$. The replication process, therefore, will introduce a replicated system made by the replicas $(3, 4)$, coupled in the same way as $(1, 2)$. Now the diagonal term gives a nontrivial contribution, and we find

$$\partial_t \tilde{\alpha}(\lambda, t) = \frac{1}{2} \langle \Delta_{12} \rangle_{\lambda, t} - \langle \Delta_{13} \rangle_{\lambda, t}, \quad (20)$$

where as before

$$\Delta_{12} \equiv S_{i_1 i_2} - \hat{S}_{i_1 i_2} \geq 0, \quad \Delta_{13} \equiv S_{i_1 i_3} - \hat{S}_{i_1 i_3} \geq 0. \quad (21)$$

The terms $\Delta_{14}, \Delta_{23}, \Delta_{24}$ do not appear explicitly, because we have exploited the $(1-2)$ and $(3-4)$ symmetries. Of course, because of the λ coupling, there is no symmetry of the type $(1-3)$, and analogous. However, at $\lambda = 0$, we have

$$\langle \Delta_{12} \rangle_{0, t} \equiv \langle \Delta_{12} \rangle_t \equiv \langle \Delta_{13} \rangle_{0, t}. \quad (22)$$

On the other hand, the λ derivative is given simply by

$$\partial_\lambda \tilde{\alpha}(\lambda, t) = \frac{1}{2} \langle \Delta_{12} \rangle_{\lambda, t}. \quad (23)$$

Introduce now a velocity field $v(\lambda, t)$, and trajectories $t \rightarrow \lambda(t)$, such that $\lambda(0) = \lambda_0$, and $\dot{\lambda}(t) = v(\lambda(t), t)$. Then we have for the total t derivative

$$\frac{d}{dt}\tilde{\alpha}(\lambda(t), t) = \frac{1}{2}(1 + v(\lambda(t), t))\langle\Delta_{12}\rangle_{\lambda,t} - \langle\Delta_{13}\rangle_{\lambda,t}. \quad (24)$$

In our previous paper [6], we made the very simple choice $v(\lambda, t) = -1$, so that $\lambda(t) = \lambda_0 - t$. In this case

$$\frac{d}{dt}\tilde{\alpha}(\lambda(t), t) = -\langle\Delta_{13}\rangle_{\lambda,t} \leq 0. \quad (25)$$

Therefore

$$\tilde{\alpha}(\lambda(t), t) \leq \tilde{\alpha}(\lambda_0, 0). \quad (26)$$

A more efficient way is to go along the level lines of $\tilde{\alpha}$, with the natural choice

$$v(\lambda, t) = -1 + 2\frac{\langle\Delta_{13}\rangle_{\lambda,t}}{\langle\Delta_{12}\rangle_{\lambda,t}}, \quad (27)$$

so that

$$\frac{d}{dt}\tilde{\alpha}(\lambda(t), t) = 0, \quad \tilde{\alpha}(\lambda(t), t) \equiv \tilde{\alpha}(\lambda_0, 0). \quad (28)$$

This choice will play a central role in our treatment. Notice that $v(0, t) \equiv 1$. Let us now derive the lower bound. From the definition (19) of $\tilde{\alpha}$, by applying Jensen inequality, we easily get

$$\begin{aligned} \tilde{\alpha}(\lambda(t), t) &\equiv \alpha(t) + \frac{1}{2}E \log \Omega(\exp(\lambda(t)\Delta_{12})) \\ &\geq \alpha(t) + \frac{\lambda(t)}{2}\langle\Delta_{12}\rangle_t. \end{aligned} \quad (29)$$

By recalling (16) and (26) or (28), we obtain

$$\lambda(t)\frac{d}{dt}(\alpha(0) - \alpha(t)) \leq \tilde{\alpha}(\lambda_0, 0) - \alpha(t). \quad (30)$$

Provided the condition $\lambda(t') > 0$ holds for $0 \leq t' \leq t$, we can easily integrate (30), and derive the basic lower bound for $\alpha(t)$

$$0 \leq \alpha(0) - \alpha(t) \leq (\tilde{\alpha}(\lambda_0, 0) - \alpha(0))(\exp(\int_0^t \frac{dt'}{\lambda(t')}) - 1). \quad (31)$$

In the following we will make extensive use of this lower bound.

3 The mean field spin glass model in the replica symmetric region

The generic configuration of the mean field spin glass model is defined through Ising spin variables $\sigma_i = \pm 1$, attached to each site $i = 1, 2, \dots, N$.

The Hamiltonian of the model has a random character, and can be defined as follows. First of all let us introduce a family of centered Gaussian random variables $\mathcal{K}(\sigma)$, indexed by the configurations σ , and characterized by the covariances

$$E(\mathcal{K}(\sigma)\mathcal{K}(\sigma')) = q^2(\sigma, \sigma'), \quad (32)$$

where $q(\sigma, \sigma')$ are the overlaps between two generic configurations, defined by

$$q(\sigma, \sigma') = N^{-1} \sum_i \sigma_i \sigma'_i, \quad (33)$$

with the obvious bounds $-1 \leq q(\sigma, \sigma') \leq 1$, and the normalization $q(\sigma, \sigma) = 1$.

At a generic inverse temperature β , for an external magnetic field of strength h , the (random) partition function is defined in the form

$$Z_N(\beta, h; \mathcal{K}) = \sum_{\sigma_1 \dots \sigma_N} \exp(\beta \sqrt{\frac{N}{2}} \mathcal{K}(\sigma)) \exp(\beta h \sum_i \sigma_i), \quad (34)$$

which will be the starting point of our treatment.

By now standard methods [13] assure the existence of the infinite volume limit in the form

$$\lim_{N \rightarrow \infty} N^{-1} E \log Z_N(\beta, h; \mathcal{K}) \equiv \sup_N N^{-1} E \log Z_N(\beta, h; \mathcal{K}) \equiv \alpha(\beta, h). \quad (35)$$

Now we introduce the Sherrington-Kirkpatrick replica symmetric solution, by following the methods explained in [8].

For a generic trial parameter \bar{q} , and an interpolating $0 \leq t \leq 1$, define now

$$\begin{aligned} \alpha_N(t) &\equiv N^{-1} E \log \sum_{\sigma_1 \dots \sigma_N} \exp(\beta \sqrt{t} \sqrt{\frac{N}{2}} \mathcal{K}(\sigma)) \exp(\beta \sum_i (h + \sqrt{1-t} \sqrt{\bar{q}} J_i) \sigma_i) \\ &+ \frac{1}{2} \beta^2 (\bar{q} - 1)^2 (1 - t), \end{aligned} \quad (36)$$

where we have introduced independent unit Gaussian fields J_i , acting as magnetic fields on each site. Now E denotes average over the \mathcal{K} 's and the J_i 's.

For $t = 1$ we have

$$\alpha_N(1) \equiv N^{-1} E \log Z_N(\beta, h; \mathcal{K}), \quad (37)$$

while for $t = 0$, the *Boltzmannfaktor* factorizes, the sum over σ 's can be performed explicitly, and we obtain the replica symmetric trial function

$$\alpha_N(0) \equiv \log 2 + \int \log \cosh \beta(h + \sqrt{\bar{q}}z) d\mu(z) + \frac{1}{2} \beta^2 (\bar{q} - 1)^2. \quad (38)$$

Here $d\mu(z)$ is the unit Gaussian measure. Of course, $\alpha_N(0)$ does not depend on N , and will be written as $\alpha(0)$ in the following.

By the same procedure as in Section 2, we can easily calculate the t derivative in the form

$$\frac{d}{dt} \alpha_N(t) = -\frac{\beta^2}{4} \langle (q_{12} - \bar{q})^2 \rangle_t \leq 0. \quad (39)$$

Clearly we have the uniform bounds

$$\alpha_N(t) \leq \alpha(0), \quad \alpha_N(1) \leq \alpha(0). \quad (40)$$

In view of these bounds, since \bar{q} is a free parameter, it is convenient to take into (38) the value that minimizes $\alpha(0)$. This value $\bar{q}(\beta, h)$ is uniquely defined, as shown in [8], and satisfies the implicit equation

$$\bar{q}(\beta, h) = \int \tanh^2(\beta(h + \sqrt{\bar{q}(\beta, h)}z)) d\mu(z). \quad (41)$$

It turns out that $\bar{q}(\beta, h) > 0$, unless $h = 0$ and $\beta \leq 1$ (the annealed region), where $\bar{q}(\beta, 0) = 0$. From now on, we assume that this minimizing value $\bar{q}(\beta, h)$ has been chosen in all our formulae for \bar{q} . In this case $\alpha(0)$ in (38) gives the replica symmetric Sherrington-Kirkpatrick value $\alpha_{SK}(\beta, h) \equiv \alpha(0)$, and we have

$$\frac{d}{dt} (\alpha_{SK}(\beta, h) - \alpha_N(t)) = \frac{\beta^2}{4} \langle (q_{12} - \bar{q})^2 \rangle_t \geq 0. \quad (42)$$

Now the bounds in (40) are given in the optimal form

$$\alpha_N(t) \leq \alpha_{SK}(\beta, h), \quad \alpha_N(1) \leq \alpha_{SK}(\beta, h). \quad (43)$$

Let us now introduce the parameter $y_0 \equiv y_0(\beta, h)$ defined by

$$y_0(\beta, h) \equiv \int \cosh^{-4}(\beta(h + \sqrt{\bar{q}(\beta, h)}z)) d\mu(z). \quad (44)$$

On the (β, h) plane, $\beta \geq 0$, $h \geq 0$, define the SK region, the AT line and the P region as the regions where $\beta^2 y_0(\beta, h) \leq 1$, $\beta^2 y_0(\beta, h) = 1$ and $\beta^2 y_0(\beta, h) \geq 1$, respectively.

The main result of this paper is contained in the following theorem.

Theorem 1. *In the replica symmetric region SK, where $\beta^2 y_0(\beta, h) \leq 1$, the Sherrington-Kirkpatrick value gives the right free energy in the infinite volume limit, i.e.*

$$\lim_{N \rightarrow \infty} N^{-1} E \log Z_N(\beta, h; \mathcal{K}) \equiv \alpha(\beta, h) \equiv \alpha_{SK}(\beta, h) \equiv \inf_{\bar{q}} (\log 2 + \int \log \cosh(\beta(h + \sqrt{\bar{q}}z)) d\mu(z) + \frac{1}{2} \beta^2 (\bar{q} - 1)^2). \quad (45)$$

Let us now consider the interpolation function $\alpha_N(t)$ as defined in (36). The methods explained in [6] show that there exists $\bar{t}(\beta, h)$, $0 < \bar{t}(\beta, h) \leq 1$, such that for any $0 \leq t \leq \bar{t}(\beta, h)$, we have in the limit

$$\lim_{N \rightarrow \infty} \alpha_N(t) = \alpha_{SK}(\beta, h). \quad (46)$$

We assume that $\bar{t}(\beta, h)$ is the largest value with this property, and call the interval $0 \leq t \leq \bar{t}(\beta, h)$ the replica symmetric interval for t . Notice that this interval is always nontrivial, since $\bar{t}(\beta, h)$ is always larger than zero.

The following Theorem gives a full characterization of $\bar{t}(\beta, h)$.

Theorem 2. *In the replica symmetric region SK, where $\beta^2 y_0(\beta, h) \leq 1$, $\bar{t}(\beta, h)$ takes its maximum value $\bar{t}(\beta, h) = 1$. In the broken replica region, where $\beta^2 y_0(\beta, h) > 1$, the $\bar{t}(\beta, h)$ takes the value*

$$\bar{t}(\beta, h) = \beta^{-2} y_0(\beta, h)^{-1}. \quad (47)$$

Clearly, Theorem 1 follows from the first part of Theorem 2, because we can take $t = 1$ in (46) and recall (37).

On the other hand, the first part of Theorem 2 follows from the second part, by taking the limit $\beta^2 y_0 \rightarrow 1$. The argument by Toninelli [14], based on the broken replica symmetry bounds of [4], shows that, in the broken replica symmetry region, the parameter $\bar{t}(\beta, h)$ can not be greater than $\beta^{-2} y_0(\beta, h)^{-1}$. Therefore, in order to prove Theorem 2, it is enough to show that $\bar{t}(\beta, h)$ can not be less than $\beta^{-2} y_0(\beta, h)^{-1}$. This will be our objective in the following.

4 Quadratic coupling of replicas

Now we couple two replicated systems, with spin variables $\sigma^{(1)}, \sigma^{(2)}$, in the same spirit as in [6] and in Section 2. In analogy with (19) the auxiliary function $\tilde{\alpha}_N(\lambda, t)$ is now defined, for $\lambda \geq 0$ and $0 \leq t \leq 1$, as

$$\begin{aligned} \tilde{\alpha}_N(\lambda, t) &\equiv \frac{1}{2N} E \log \sum_{\sigma^{(1)} \sigma^{(2)}} \exp(B_1) \exp(B_2) \exp\left(\frac{1}{2} \lambda \beta^2 N (q_{12} - \bar{q}(\beta, h))^2\right) \\ &+ \frac{1}{2} \beta^2 (\bar{q} - 1)^2 (1 - t), \end{aligned} \quad (48)$$

where B_1 and B_2 are the two replicas of the quantities appearing in the *Boltzmannfaktor* of $\alpha_N(t)$ in (36).

The function $\tilde{\alpha}_N(\lambda, t)$ is convex increasing in λ for each value of t , and we have

$$\tilde{\alpha}_N(0, t) \equiv \alpha_N(t), \quad \tilde{\alpha}_N(\lambda, t) \geq \alpha_N(t). \quad (49)$$

The t and λ partial derivatives are easily calculated in analogy with (20) and (23)

$$\partial_t \tilde{\alpha}_N(\lambda, t) = \frac{\beta^2}{4} \langle (q_{12} - \bar{q})^2 \rangle_{\lambda, t} - \frac{\beta^2}{2} \langle (q_{13} - \bar{q})^2 \rangle_{\lambda, t}, \quad (50)$$

$$\partial_\lambda \tilde{\alpha}_N(\lambda, t) = \frac{\beta^2}{4} \langle (q_{12} - \bar{q})^2 \rangle_{\lambda, t}. \quad (51)$$

Let us go to the infinite volume limit $N \rightarrow \infty$, by exploiting for example the methods outlined in [15], and call $\tilde{\alpha}(\lambda, t)$ the resulting limit of $\tilde{\alpha}_N(\lambda, t)$. For any $0 \leq t \leq \bar{t}(\beta, h)$, let $\bar{\lambda}(t)$ be the highest value of λ for which $\tilde{\alpha}(\lambda, t) = \alpha_{SK}(\beta, h)$. We call the region $0 \leq t \leq \bar{t}(\beta, h)$, $0 \leq \lambda \leq \bar{\lambda}(t)$ the replica symmetric SK region in the (λ, t) plane.

In [6], the value of $\bar{\lambda}(0)$ was calculated explicitly through a variational principle. In particular, in the case $h = 0$, $\bar{q} = 0$, it turns out that $\bar{\lambda}(0) = \beta^{-2}$. Moreover, it was shown that the triangular, or trapezoidal, region $0 \leq t \leq \min(1, \bar{\lambda}(0))$, $0 \leq \lambda \leq \bar{\lambda}(0) - t$, is contained into the replica symmetric region. Another interesting property of the replica symmetric region is the following. Assume that (λ_1, t_1) is in the SK region, then all points $t_1 \leq t \leq \min(1, t_1 + \lambda_1)$, $0 \leq \lambda \leq \lambda_1 - t + t_1$, are in the SK region. All these properties follow, according to [6], from the bound (31), adapted to the $\alpha_N(t)$ defined in (36), with the simple choice $v(\lambda, t) \equiv -1$, by taking the infinite volume limit.

A very important consequence is the following

Theorem 3. *In the case where $\bar{t}(\beta, h) < 1$, the following bound necessarily holds for $\bar{\lambda}(t)$*

$$\bar{\lambda}(t) \leq \bar{t}(\beta, h) - t, \quad (52)$$

for any $0 \leq t \leq \bar{t}(\beta, h)$.

Therefore, if $\bar{t}(\beta, h) < 1$, the replica symmetric region must shrink to the single point $\bar{\lambda}(\bar{t}(\beta, h)) \equiv 0$ at the end point $\bar{t}(\beta, h)$. The proof is straightforward. Should (52) fail, there would be a point t_1 where $\bar{\lambda}(t_1) > \bar{t}(\beta, h) - t_1$. But in this case, by the general properties pointed before, it would be possible to enlarge the replica symmetric region beyond $\bar{t}(\beta, h)$, which is clearly impossible due to the very definition of $\bar{t}(\beta, h)$.

There are two important properties of the replica symmetric region that we immediately point out.

Introduce, for generic replicas s, s' , the rescaled overlaps defined as the random variables

$$\xi_{ss'} \equiv \sqrt{N}(q_{ss'} - \bar{q}(\beta, h)), \quad (53)$$

distributed according to the $\langle \dots \rangle_{\lambda, t}$ state. Then we have

Theorem 4. *For any (λ, t) in the replica symmetric region, in the infinite volume limit, the ξ become centered Gaussian random variables, whose covariances can be explicitly calculated. These covariances may become singular only along the boundary line $\bar{\lambda}$, or at the boundary point $\bar{t}(\beta, h)$, according to the various situations analyzed in the following.*

The proof is very simple. In fact, it is a simple adaptation of the same procedure exploited in [7], and [3], for the case where $\lambda = 0$. The effective calculation of the covariances exploits methods introduced in [8], [7], and [3]. We sketch the extension to the case where $\lambda > 0$ in the next section 5.

In the replica symmetric region, we have not only the convergence of the $\tilde{\alpha}_N(\lambda, t)$ auxiliary function to $\alpha_{SK}(\beta, h)$, but also a stronger statement.

Theorem 5. *For any (λ, t) in the replica symmetric region, with the possible exclusion of the end point $\bar{t}(\beta, h)$, the infinite volume limit*

$$\lim_{N \rightarrow \infty} N(\tilde{\alpha}_N(\lambda, t) - \alpha_{SK}(\beta, h)) \equiv \phi(\lambda, t) \quad (54)$$

is well defined and can be explicitly given in terms of integrals of fluctuations as follows

$$\phi(\lambda, t) = -\frac{\beta^2}{4} \int_0^t A(t') dt' + \frac{\beta^2}{4} \int_0^\lambda A(\lambda', t) d\lambda', \quad (55)$$

where

$$A(t') \equiv \lim_{N \rightarrow \infty} \langle \xi_{12}^2 \rangle_{t'}, \quad A(\lambda', t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{12}^2 \rangle_{\lambda', t}. \quad (56)$$

The function ϕ could take the value ∞ at the boundary line $\bar{\lambda}$, according to the various situations analyzed in the following.

The proof is very simple. In fact, we recall that $\tilde{\alpha}_N(0, t) \equiv \alpha_N(t)$, and write

$$N(\tilde{\alpha}_N(\lambda, t) - \alpha_{SK}(\beta, h)) \equiv -N(\alpha_{SK}(\beta, h) - \alpha_N(t)) + N(\tilde{\alpha}_N(\lambda, t) - \tilde{\alpha}_N(0, t)). \quad (57)$$

Then we exploit (42) and (51). By integration we get

$$N(\alpha_{SK}(\beta, h) - \alpha_N(t)) = \frac{\beta^2}{4} \int_0^t \langle \xi_{12}^2 \rangle_{t'} dt', \quad (58)$$

$$N(\tilde{\alpha}_N(\lambda, t) - \tilde{\alpha}_N(0, t)) = \frac{\beta^2}{4} \int_0^\lambda \langle \xi_{12}^2 \rangle_{\lambda', t} d\lambda'. \quad (59)$$

By taking the infinite volume limit we get (55), and Theorem 5 is proven.

It is also convenient to introduce the level lines of the function ϕ . In fact, we follow the procedure exploited in section 2, by starting with the level lines for the the difference $\tilde{\alpha}_N(\lambda, t) - \alpha_{SK}(\beta, h)$. It is immediately recognized, as in section 2, that, for finite N , the corresponding velocity field is given by

$$v(\lambda, t) = -1 + 2 \frac{\langle (q_{13} - \bar{q})^2 \rangle_{\lambda, t}}{\langle (q_{12} - \bar{q})^2 \rangle_{\lambda, t}} \equiv -1 + 2 \frac{\langle \xi_{13}^2 \rangle_{\lambda, t}}{\langle \xi_{12}^2 \rangle_{\lambda, t}}. \quad (60)$$

In the replica symmetric region, fluctuations can be completely controlled. Therefore, in the infinite volume limit, we have that the velocity field of the level lines for the limiting ϕ function are given explicitly by

$$v(\lambda, t) = -1 + 2 \frac{D(\lambda, t)}{A(\lambda, t)}, \quad (61)$$

where $A(\lambda, t)$ is defined in (56), and

$$D(\lambda, t) = \lim_{N \rightarrow \infty} \langle \xi_{13}^2 \rangle_{\lambda, t}. \quad (62)$$

Notice that $A(\lambda, t)$ and $D(\lambda, t)$ are different in general, because the λ coupling destroys the $(2-3)$ symmetry. They are explicitly calculated by following the method outlined in section 5. Of course, at $\lambda = 0$, we have $A(0, t) \equiv D(0, t) \equiv A(t)$. In particular $v(0, t) \equiv 1$, for any $t < \bar{t}(\beta, h)$, where $A(t)$ is surely not singular.

Let us give a simple example. We consider the annealed case, where $h = 0$, $\bar{q} = 0$. Here everything can be explicitly calculated. In particular, the methods of [6] give $\bar{t}(\beta) = 1$ if $\beta \leq 1$, and $\bar{t}(\beta) = \beta^{-2}$ if $\beta > 1$. The annealed region is given by $0 \leq t \leq \bar{t}(\beta)$, $0 \leq \lambda \leq \bar{\lambda}(t)$, where $\bar{\lambda}(t) = \beta^{-2} - t$. In the annealed region, by following the methods outlined in section 5, we find that the variances are given by

$$A(t) = 1/(1 - t\beta^2), \quad D(\lambda, t) \equiv A(t), \quad A(\lambda, t) = 1/(1 - (\lambda + t)\beta^2). \quad (63)$$

Notice that $A(\lambda, t)$ becomes infinite on the $\bar{\lambda}$ line, where $v(\bar{\lambda}(t), t) = -1$. In general we have for the velocity field, from (61) and (63),

$$v(\lambda, t) = -1 + 2 \frac{D(\lambda, t)}{A(\lambda, t)} \equiv 1 + 2 \frac{\lambda\beta^2}{1 - t\beta^2} \quad (64)$$

The level lines are immediately found by integration of the streaming equation $\dot{\lambda}(t) = v(\lambda(t), t)$. Let λ_0 be a parameter with the condition $\lambda_0 \leq \bar{\lambda}_0 \equiv \beta^{-2}$. Then the family of level lines is given by

$$\lambda(t) = (\beta^{-2} - t) - (\beta^{-2} - \lambda_0)(1 - \beta^2 t)^2, \quad (65)$$

where in any case it must be $0 \leq t \leq 1$ and $\lambda(t) \geq 0$. If $\lambda_0 \geq 0$, then the trajectories start from λ_0 at $t = 0$. If $\lambda_0 < 0$ then the trajectories start from $\lambda = 0$ at $t = \lambda_0/(1 - \beta^2 \lambda_0)$. The value of the function ϕ , as defined in (54), can be easily found from (55) in the form

$$\phi(\lambda, t) = \frac{1}{4} \log \frac{1}{1 - \lambda_0 \beta^2}, \quad (66)$$

where λ_0 is the parameter characterizing the trajectory passing through (λ, t) .

In this simple completely solvable annealed case, we can point out the key phenomenon appearing on the (λ, t) plane. Assume $\beta > 1$. Then we know $\bar{t}(\beta) \equiv \beta^{-2} < 1$. The annealed region is the triangle $0 \leq t \leq \bar{t}(\beta)$, $0 \leq \lambda \leq \beta^{-2} - t$. At the end point $t = \bar{t}(\beta)$, $\lambda = 0$, the variance $A(t)$ becomes infinite, the function $\phi(0, \bar{t}(\beta))$ is not well defined, and moreover all trajectories have there an osculating focus with common slope $v(0, \bar{t}(\beta)) = -1$.

We will show in the following that this kind of picture holds qualitatively also in the general replica symmetric case.

5 Calculation of overlap fluctuations

Overlap fluctuations have been calculated in the infinite volume limit in [8] and [7], see also [3].

Firstly, we consider the case $\lambda = 0$. The initial conditions at $t = 0$ follow easily from the central limit theorem. In fact we have

$$A_0 \equiv \lim_{N \rightarrow \infty} \langle \xi_{12}^2 \rangle_0 \equiv 1 - \bar{q}^2, \quad B_0 \equiv \lim_{N \rightarrow \infty} \langle \xi_{12} \xi_{13} \rangle_0 \equiv \bar{q} - \bar{q}^2, \quad (67)$$

$$C_0 \equiv \lim_{N \rightarrow \infty} \langle \xi_{12} \xi_{34} \rangle_0 \equiv \int \tanh^4(\beta(h + \sqrt{\bar{q}(\beta, h)}z)) d\mu(z) - \bar{q}^2. \quad (68)$$

It is also convenient to define

$$y_0 \equiv A_0 - 2B_0 + C_0, \quad z_0 \equiv A_0 - 4B_0 + 3C_0. \quad (69)$$

It is immediately recognized that y_0 has the value already given in (44). Moreover, the bound $z_0 \leq y_0$ holds.

In the replica symmetric interval, in the infinite volume limit, the ξ become centered Gaussian random variables, with explicitly given covariances (see [8], [7], and [3]) according to

$$A(t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{12}^2 \rangle_t = 3y(t) - 2z(t) + w(t), \quad (70)$$

$$B(t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{12} \xi_{13} \rangle_t = \frac{3}{2}y(t) - \frac{3}{2}z(t) + w(t), \quad (71)$$

$$C(t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{12} \xi_{34} \rangle_t = y(t) - z(t) + w(t), \quad (72)$$

where

$$y(t) = \frac{y_0}{1 - t\beta^2 y_0}, \quad z(t) = \frac{z_0}{1 - t\beta^2 z_0}, \quad w(t) = \frac{3C_0 - 2B_0}{(1 - t\beta^2 z_0)^2}. \quad (73)$$

In order to calculate the covariances for $\lambda \geq 0$ we follow a very simple strategy. For any t in the replica symmetric interval, we take $A(t), B(t), C(t)$ as initial conditions, and consider the streaming equations in λ . These are easily solved in the infinite volume limit. There is a mild complication. Since the symmetry between replicas is explicitly broken by the λ couplings, there are more basic covariances to be taken into account. Let us consider some examples.

Define, in agreement with (56),

$$A(\lambda, t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{12}^2 \rangle_{\lambda, t}, \quad C(\lambda, t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{12} \xi_{34} \rangle_{\lambda, t}, \quad (74)$$

with the initial conditions $A(0, t) = A(t)$, $C(0, t) = C(t)$. At fixed t , the streaming equations are

$$\frac{d}{d\lambda} \langle \xi_{12}^2 \rangle_{\lambda, t} = \frac{\beta^2}{2} \langle \xi_{12}^2 (\xi_{12}^2 - \xi_{34}^2) \rangle_{\lambda, t}, \quad (75)$$

$$\frac{d}{d\lambda} \langle \xi_{12} \xi_{34} \rangle_{\lambda, t} = \frac{\beta^2}{2} \langle \xi_{12} \xi_{34} (\xi_{12}^2 + \xi_{34}^2 - \xi_{56}^2 - \xi_{78}^2) \rangle_{\lambda, t}. \quad (76)$$

Now we go to the infinite volume limit, take into account the Gaussian nature of the resulting fluctuations, and exploit the partial symmetry among replicas. We find the equations

$$A'(\lambda, t) = A(\lambda, t)^2 - C(\lambda, t)^2, \quad C'(\lambda, t) = 2A(\lambda, t)C(\lambda, t) - 2C(\lambda, t)^2, \quad (77)$$

where the prime denotes the partial derivative with respect to the rescaled parameter $\lambda\beta^2$. These equations, with the given initial conditions, are easily integrated in the form

$$A(\lambda, t) = \frac{A(t) - C(t)}{1 - \lambda\beta^2(A(t) - C(t))} + C(\lambda, t), \quad (78)$$

$$C(\lambda, t) = \frac{C(t)}{(1 - \lambda\beta^2(A(t) - C(t)))^2}. \quad (79)$$

Analogous procedure can be followed for the other covariances. For example, for

$$B(\lambda, t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{12} \xi_{13} \rangle_{\lambda, t}, \quad D(\lambda, t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{13}^2 \rangle_{\lambda, t}, \quad E(\lambda, t) \equiv \lim_{N \rightarrow \infty} \langle \xi_{13} \xi_{56} \rangle_{\lambda, t}, \quad (80)$$

we find the equations

$$B' = AB + BC - 2CE, \quad D' = 2B^2 - 2E^2, \quad E' = 2BC + AE - 3CE, \quad (81)$$

which are easily solved by taking into account the initial conditions $B(0, t) = B(t)$, $D(0, t) = A(t)$, $E(0, t) = C(t)$.

We end this section with the following qualitative estimates, which are useful for the investigation of the properties of the level lines introduced in the previous section.

Theorem 6. *The following estimates hold*

$$A(\lambda, t) \leq \frac{A(t)}{1 - \lambda\beta^2 A(t)}, \quad D(\lambda, t) \geq A(t), \quad v(\lambda, t) \geq 1 - 2\lambda\beta^2 A(t). \quad (82)$$

The proof is simple. By neglecting the C^2 term in the first equation in (77), we obtain $A' \leq A^2$, which gives the first inequality in (82) upon integration. On the other hand, by exploiting (81), we find

$$(B - E)' = (B - E)(A - C). \quad (83)$$

In general $A \geq C$ by Schwarz inequality, and $B \geq E$ at $\lambda = 0$, because here it reduces to the obvious $B(t) \geq C(t)$. Therefore $B \geq E$ for any λ . But then, from (81), we derive $D' \geq 0$, and establish $D(\lambda, t) \geq A(t)$. Finally, the lower bound for v comes from the previous bounds, and the definition (61).

6 The replica symmetric region extends up to the Almeida-Thouless line

Now we are ready to prove Theorem 2. Assume that $\beta^2 y_0(\beta, h) > 1$. We show that $\bar{t}(\beta, h)$ must have exactly the value $\bar{t}(\beta, h) = \beta^{-2} y_0(\beta, h)^{-1}$. By Toninelli argument in [14], we already know $\bar{t}(\beta, h) \leq \beta^{-2} y_0(\beta, h)^{-1}$. Suppose $\bar{t}(\beta, h) < \beta^{-2} y_0(\beta, h)^{-1}$. Then the variance $A(t)$ would stay finite up to the point $t = \bar{t}(\beta, h)$ included, as the explicit solution in (70) shows (recall that $z_0 \leq y_0$). As a consequence, the function ϕ , defined in (54), would be well defined at the point $(\lambda = 0, t = \bar{t}(\beta, h))$, due to the integral representation (55) taken at $\lambda = 0$. Here we find an immediate contradiction. In fact, due to the condition (52), in the point $(\lambda = 0, t = \bar{t}(\beta, h))$ all level lines should converge. But each of them carries its own value for the function ϕ , depending on the initial condition λ_0 . For example, along the trajectory starting from $\lambda = 0$ at $t = 0$, we must have $\phi = 0$. As a general fact, the

level lines can not focus on a point where ϕ is well defined. We conclude that $\bar{t}(\beta, h) < \beta^{-2}y_0(\beta, h)^{-1}$ is impossible. We must have $\bar{t}(\beta, h) = \beta^{-2}y_0(\beta, h)^{-1}$. Therefore, Theorem 2 is proven, together with all results of this paper.

We end this section with an argument showing the unwillingness of the level lines to focus in a point where $A(t)$ is not singular. Let us recall that we have proven the bound $v(\lambda, t) \geq 1 - 2\lambda\beta^2 A(t)$ (see (82)), for the velocity field of the level lines. Now assume that $A(t)$ is finite along the whole replica symmetric interval, so that $A(t) \leq a$, for some positive constant a . Then we would have $v(\lambda, t) \geq 1 - 2\lambda\beta^2 a$. This can be easily integrated. For example, we would find that the level line starting from $\lambda = 0$ at $t = 0$, should stay above the line

$$\bar{\lambda}(t) = (1 - \exp(-2t\beta^2 a))/2\beta^2 a, \quad (84)$$

which is clearly incompatible with (52).

In conclusion, our general argument shows that, on the (β, h) plane, the boundary between the replica symmetric region and the broken replica symmetry region is characterized by a singularity in the rescaled overlap fluctuations, exactly as happens in the well known case of vanishing magnetic field. Therefore the transition is exactly at the Almeida-Thouless line.

It is interesting to compare our results with the characterization of the transition line given by the very clever method of Talagrand [3]. Talagrand exploits broken replica symmetry bounds for a system made by two coupled replicas. As a result, the transition line is given in terms of the onset of the instability of the replica symmetric solution in the frame of a one step breaking of the replica symmetry in the Parisi representation. Therefore, the problem of comparing the Talagrand transition line with the Almeida-Thouless line is open. In our method, we do not employ any information coming from the breaking of the replica symmetry, but exploit only the peculiar properties of the quadratic replica coupling. It is also important to remark that Talagrand method has been extended to give the proof of the validity of the Parisi representation for the free energy in the infinite volume limit [5]. In next section we deal with the problem of extending our method to more general situations.

7 Conclusion and outlook for future developments

We have seen that the control of the rescaled overlap fluctuations and the peculiar properties of the quadratic replica coupling allow to show that replica symmetry holds up to the Almeida-Thouless line. No information coming

from the pattern of the spontaneous replica symmetry breaking phenomenon is necessary.

In its very essence, this method allows to isolate an important parameter ruling the phenomenon. The replica symmetric representation holds, provided the parameter $\beta^2(A_0 - 2B_0 + C_0)$ is less or equal to 1. Therefore, the ruling parameter is given in terms of fluctuations at zero coupling!

It is important to remark that the flow equations for fluctuations do not depend on the particularly chosen representation. The solutions do depend on the representation, only through the initial conditions (A_0, B_0, C_0) .

Therefore, it is possible to extend our method also to other representations. In particular, in the frame of the Aizenman-Sims-Starr extended variational principle [16], we can take the representation based on Derrida-Ruelle-Parisi probability cascades (see for example [17]). Here sum rules also hold, and we can consider a generalized quadratic replica coupling between replicas. We plan to report on these problems in a future publication.

Acknowledgments

We gratefully acknowledge useful conversations with Dmitry Panchenko, Michel Talagrand, and Fabio Toninelli. Their constructive criticism was very useful in the improvement of the presentation of the results contained in this paper. The strategy exploited in this paper grew out from a systematic exploration of comparison and interpolation methods, developed in collaboration with Fabio Toninelli, and Luca De Sanctis.

This work was supported in part by MIUR (Italian Minister of Instruction, University and Research), and by INFN (Italian National Institute for Nuclear Physics).

References

- [1] D. Sherrington and S. Kirkpatrick, *Solvable Model of a Spin-Glass*, Phys. Rev. Lett. **35**, 1792-1796 (1975).
- [2] M. Mézard, G. Parisi and M. A. Virasoro, *Spin glass theory and beyond*, World Scientific, Singapore, 1987.
- [3] M. Talagrand, *Spin glasses: a challenge for mathematicians. Mean field models and cavity method*, Springer-Verlag, Berlin (2003).
- [4] F. Guerra, *Broken replica symmetry bounds in the mean field spin glass model*, Commun. Math. Phys. **233**, 1-12 (2003).

- [5] M. Talagrand, *The Parisi formula*, Annals of Mathematics, **163**, 221-263 (2006).
- [6] F. Guerra and F.L. Toninelli, *Quadratic replica coupling in in the Sherrington-Kirkpatrick mean field spin glass model*, Journal of Mathematical Physics **43**, 3704-3716 (2002).
- [7] F. Guerra and F.L. Toninelli, *Central limit theorem for fluctuations in the high temperature region of the Sherrington-Kirkpatrick spin glass model*, Journal of Mathematical Physics **43**, 6224-6237 (2002).
- [8] F. Guerra, *Sum rules for the free energy in the mean field spin glass model*, Fields Institute Communications **30**, 161 (2001).
- [9] K. Joag-dev, M.D. Perlman and L.D. Pitt, *Association of normal random variables and Slepian's inequality*, Annals of Probability **11**, 451-455 (1983).
- [10] J.-P. Kahane, *Une inégalité du type Slepian and Gordon sur les processus gaussiens*, Israel J. Math. **55**, 109-110 (1986).
- [11] S. Franz and M. Leone, *Replica bounds for optimization problems and diluted spin systems*, J. Stat. Phys. **111**, 535-564 (2003).
- [12] L. De Sanctis, *Structural Approaches to Spin Glasses and Optimization Problems*, Ph.D. Thesis, Department of Mathematics, Princeton University, 2005.
- [13] F. Guerra and F. L. Toninelli, *The Thermodynamic Limit in Mean Field Spin Glass Models*, Commun. Math. Phys. **230**, 71-79 (2002).
- [14] F. L. Toninelli, *About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean-field spin glass model*, Europhy. Lett. **60**, 764-767 (2002).
- [15] F. Guerra and F. L. Toninelli, *The infinite volume limit in generalized mean field disordered models*, Markov Processes and Related Fields **9**, 195-207 (2003).
- [16] M. Aizenman, R. Sims and S. Starr, *Extended variational principle for the Sherrington-Kirkpatrick spin-glass model*, Phys. Rev. **B68**, 214403 (2003).

- [17] F. Guerra, *An introduction to mean field spin glass theory: methods and results*, Lectures given at the Les Houches Summer School, June 2005, to appear.